50p.

N64-20597
CAT. 14-CODE-1
NASA CR-53501

FINAL REPORT

VOICE/CARDIAC SOUNDS UNDER STRESS

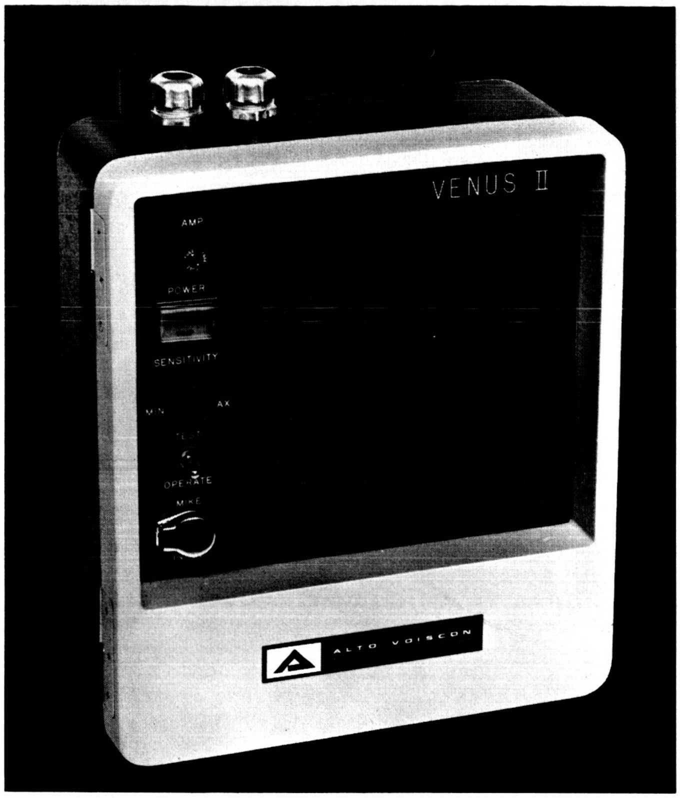
Contract NAS 2-1394

Submitted By:

Alto Scientific Company, Inc.

28 August 1963

ALTO SCIENTIFIC COMPANY



VOISCON

Voiscon (an abbreviation of "Voice Control") is a new concept in electrical switches. It responds to the voiced commands of a machine operator, meanwhile screening out other normal sounds. Here is a method of "no hands" control offering a limitless variety of industrial applications. Almost any number of machine control functions can be performed by an operator through the use of voice control, leaving his hands free for other more important work. Voiscon will hear and respond to a voiced command even in the presence of very loud surrounding shop noise, or from a considerable distance from the microphone.

VOISCON FOR PRODUCTION

★ SETTING UP large and complex machines where control can be made more accurate with both hands and voice.

★ STOPPING a machine quickly when the operator may not be near the switch to prevent damage to the machine or production process.

★ STARTING OR STOPPING one machine function with a voiced command while controlling another function manually.

VOISCON FOR SAFETY

★ STOPPING a machine quickly in an emergency situation where the danger of personal injury is imminent.

In institutional applications, Voiscon opens up an entirely new field of machine control.

VOCATIONAL SCHOOLS – Giving the instructor the capability of stopping or starting any machine in any area, regardless of where the instructor may be in the room.

REHABILITATION – Providing handicapped persons who have paralyzed or missing extremities the opportunities to operate equipment through the use of only their voices.

HOSPITALS AND THERAPY – Starting up wheel chairs, turning on radios, turning pages of books, calling nurses when infants are crying – help for the helpless.

SPECIFICATIONS

FOR ALTO SCIENTIFIC COMPANY "VOISCON"

1 Input–up to four microphones in parallel, and a "carry-on" microphone which, when plugged in, disables the other four microphones.

2 Microphone Type – Shure Model 401A, furnished with unit.

3 Ambient Noise Level – Up to 110 db.

4 Controls – Power on-off with indicator
Test-operate selector with indicator
Sensitivity
Single-double command selector

5 Output – One set of dry contacts (NO or NC) rating: 1 amp @ 115VAC or 30VDC.

6 Control Functions–Built to customer specifications and installed in a separate unit. Possible functions are Latch Stop-Manual Reset, Timed Go, Stop-Go on alternate commands, and many other functions.

7 Power Source – 115 ± 10% VAC, 60 cps, single phase.

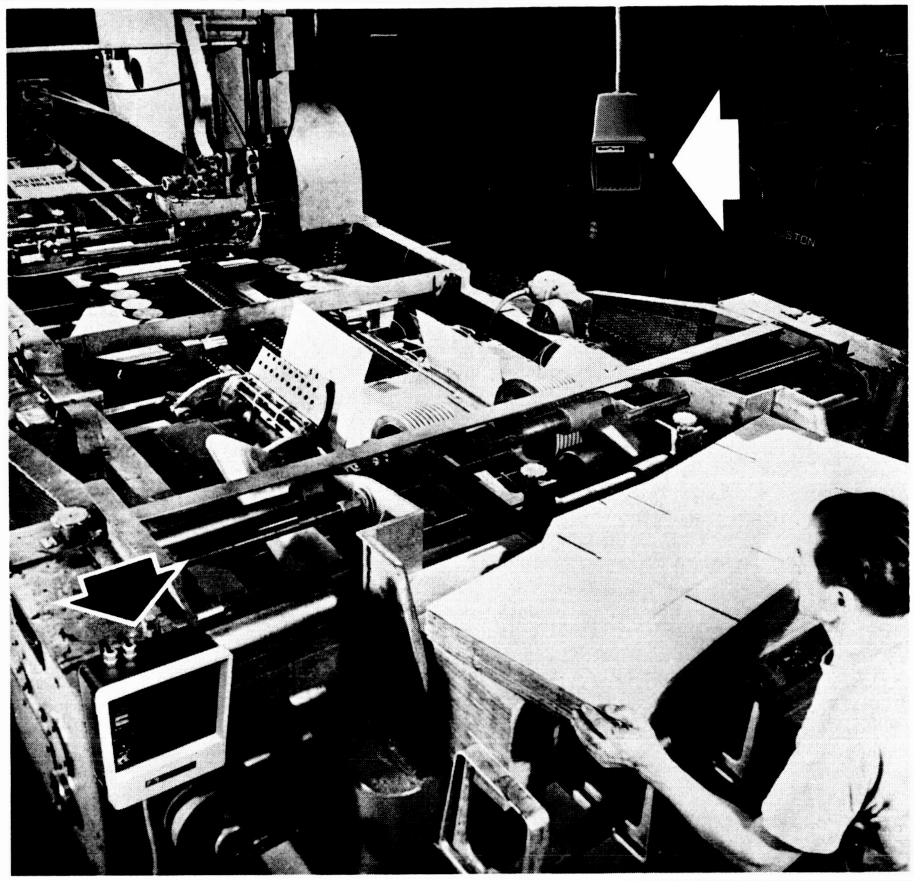
8 Size – 10" wide by 12" high by 5" deep. Weight: 15 lbs.

for field or engineering help with your machine control problems, call or write

ALTO SCIENTIFIC COMPANY
4083 TRANSPORT STREET, PALO ALTO, CALIFORNIA ■ PHONE: 321-3434

FASTER STOPS CUT DOWN MACHINE DAMAGE—LESSEN DOWN TIME

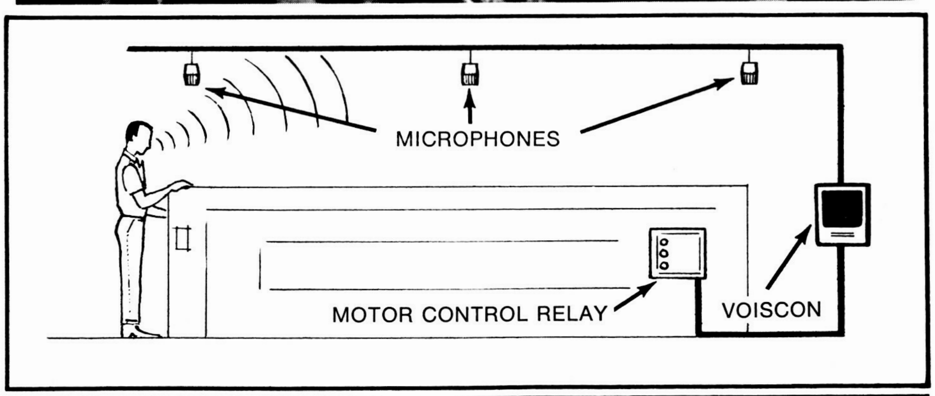
Voiscon control reduces the severity of machine jam-up... the quick stop makes it easier to clear the machine and get it back into profitable production.



Microphones (white arrow) are placed above the machine, thus letting the operator control the machine from any point—even when his hands are busy.

Voiscon switch (black arrow) is mounted on or near the machine allowing operator quick access to the controls.

VOISCON INSTALLS EASILY ON EXISTING EQUIPMENT



Microphones are simply plugged into the jacks on the bottom of the Voiscon switch. The output line of the Voiscon control is then connected to the motor control relay to interrupt the circuit on the voiced command. Any size machine can be controlled by increasing the number of microphones.

VOISCON SAVES MONEY AND IMPROVES SAFETY

The need to adjust, clean or service machines while running, requires that some type of safety guard be incorporated to protect the operator. Voiscon, in one simple control guards the **entire machine** by allowing the operator to stop it quickly with one voiced command. This quick stop may be necessary to save the machine from damage, but it can also save an operator from serious injury should he become entangled in the machine.

TYPICAL VOISCON APPLICATIONS

- **Printing presses (release pressure on cylinder)**
- **Container forming machines**
- **Painting and coating processes**
- **Saturators (for stop and go threading)**
- **Metal forming machines**
- **Rubber mixing mills**
- **Conveyors, assembly lines**

N64-20597

CAT.-14 CODE-1

NASA CR-53501

FINAL REPORT

VOICE/CARDIAC SOUNDS UNDER STRESS

Contract NAS 2-1394

Submitted By:

Alto Scientific Company, Inc.

28 August 1963

CONTENTS

REPORT

ILLUSTRATIONS

APPENDICES

A Table of Phonetic Symbols

B Description of Sigmatometer for Measurement of Fricative Competence

C Description of Voiscon as a Noise Tolerant Voice Switch

D Definition of Terms

E Chronology of Centrifuge Data Runs

1.0 Study Objective

There were two objectives for the study performed under Contract NAS 2-1394:

1) Determine if Alto speech recognition detect circuits can measure changes in a pilot's speech under varying environmental C stress.

2) Make a preliminary evaluation to determine whether peak power asymmetry measuring techniques show promise in obtaining a sophisticated phonocariogram - particularly as related to measuring change in cardiac sounds under varying G stress.

Inherent in these two objectives is the desire to develop additional techniques for determining pilot well-being in manned space flights and to do so within the practical limits dictated by vehicle environments.

2.0 Introduction

As a result of listening to voice reports made by pilots in manned centrifuges, many persons within NASA have noted that the pilot's voice "quality" and ability to form certain sounds were affected under conditions of high G stress. In addition, as a result of the effect of hydraulic displacement upon cardiac valving, some physiologists surmise that sounds emanating from the heart might also change under acceleration. Thus the purpose of this study was to determine whether these changes could be objectively measured.

3.0 Technical Background

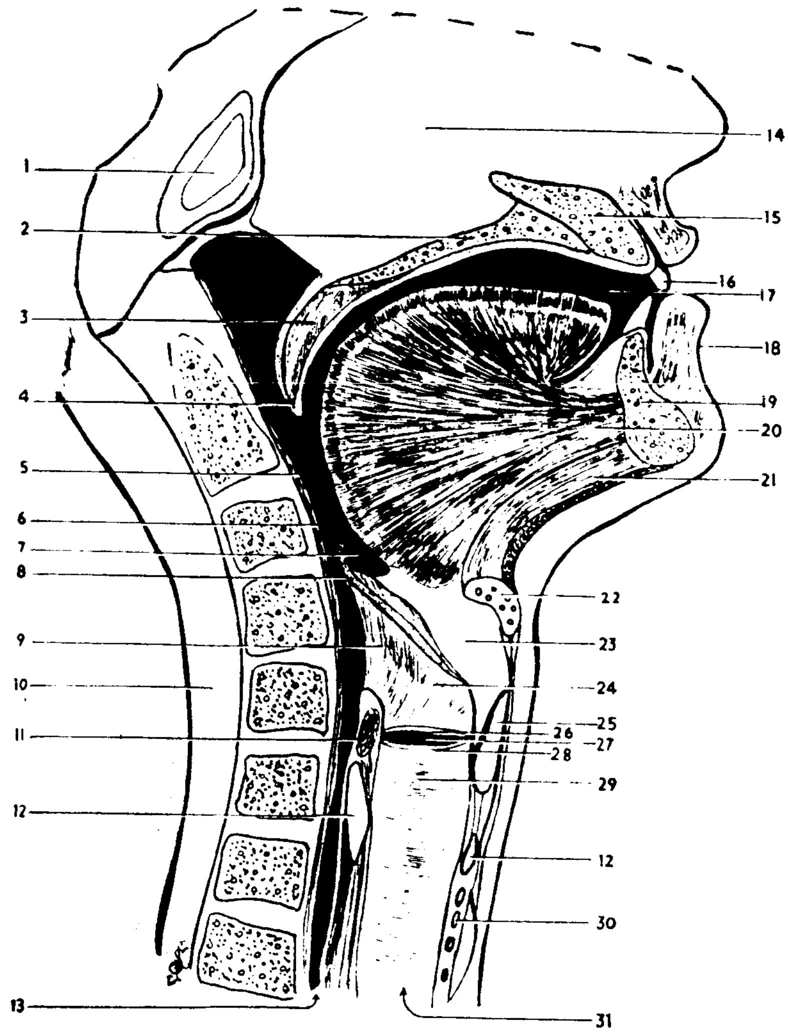
3.1 Production of Speech

Voice sounds are those sounds produced by laryngeal vibration (often referred to as vocal fold vibration, phonation, or simply voicing). Phonation is accompanied by resonance effects in the trachea, the pharynx, the oral cavity, and the nasal cavity. Speech sounds, for purposes of this report, will be considered those purposive, communicative sounds that include phonation and/or articulation (plus resonances associated with both), where articulation is the controlled interruption of the phonated or non-phonated breath stream.

In order to present examples of speech sounds, use is made of the "phonetic alphabet" (see Appendix A) wherein sounds are described by symbols. Phonetic symbols will be shown underlined in parenthesis, e.g., "food" would be shown phonetically as (<u>fud</u>). Phonetic symbols do not represent a single sound, but rather a series of slightly different sounds that include all variations which are perceived acoustically as the sound under consideration. Such sound families are referred to as "phonemes". Measurements made of speech sounds can be expected to vary within the limits of the accuracy of phonemic production.

An example of a purely phonated sound is the word "all" (<u>ɔl</u>). This sound is produced roughly as follows: (See Illustration 1) Vibration is produced as a saw-tooth waveform at the vocal folds, item 28, by the controlled outflow of a tracheal airstream. The rate of vibration, or pitch, for this sound would fall into 100 to 200 cps range for the average male adult. The oral resonation chamber transcends from a tongue depressed, open mandible position for the (<u>ɔ</u>) sound to a tongue-alveolar contact and somewhat less open mandible position for (<u>l</u>) sound. Coupling into the nasal cavity is controlled by the position of the velum, or soft palate, item 3. The sound (<u>ɔl</u>) is considered primarily non-nasal because the major fraction of total sound energy is orally rather than nasally directed. It is apparent, then, that the major resonating cavities over which control can be exercised is (a) the coupling to the nasal cavities, which are in themselves of fixed size, and (b) the size and division of the oral cavity by means of the mandible position, the tongue position, and lip position. Measurement of speech sounds would exhibit variations if G stresses caused partial or total loss of control over these cavities.

The cavity openings tend to reinforce certain frequency components present in vocal fold vibration which gives all sounds an individual "character". This character can be described by formant structure, i.e., frequency bands in which energy is concentrated during the production of a speech sound.



Median Sagittal Section Through The Head And Neck

1. Sphenoidal sinus; 2. Hard Palate; 3. Soft Palate; 4. Uvula; 5. Superior Pharyngeal Constrictor; 6. Middle Pharyngeal Constrictor; 7. Vallecula; 8. Epiglottis; 9. Aditus Laryngis; 10. Vertebral Canal; 11. Arytenoidei; 12. Cricoid Cartilage; 13. Esophagus; 14. Nasal Cavity; 15. Maxilla; 16. Teeth; 17. Oral Cavity; 18. Lip; 19. Mandible; 20. Genioglossus; 21. Geniohyoid; 22. Hyoid Bone; 23. Adipose Tissue; 24. Hyperglottal Cavity; 25. Thyroid Cartilage; 26. Superior Ventricular Folds; 27. Glotus; 28. Vocal Folds; 29. Hypoglottal Cavity; 30. Tracheal Cartilages; 31. Trachea

ILLUSTRATION 1

3.2 Classification and Competence of Speech Sounds

Speech sounds can be classified as follows: vowels, glides, and consonants. The vowels can be divided into rough degrees of mandible opening as follows: (a) closed, (b) intermediate, and (c) open. Corresponding examples are: (a) eat (it), (b) up (ʌp), and (c) all (ɔl). Glides (often called liquids or semi-vowels) are the sounds resulting from the connection of one vowel to a second vowel while maintain ing continuous voicing. Examples of the glides are the sounds (l), (r), (w) and the diphthongs. Consonants can be divided into four major groups; the stop plosives, the nasals, the continuant fricatives, and the affricates. The plosives and fricatives can be further divided into surd (voiceless) and sonnet (voiced) sounds. Examples of these sounds follow: surd stop plosive, "two" (tu), nasal "no" (no); sonnet continuant fricative, "zoo" (zu); affricate, "cats" (kæts). Competence of speech sounds generally relates to the accuracy of a phonemic production. Previous training dictates the limits of acceptability so that, in simplified terms, any speech is in a delicate balance of continual correction by the comparison of actual production to the "stored" information concerning correct production via the speakers aural (hearing) mechanism. Competence as determined by perception of a single speech sound is without degree; either an acceptable phoneme is produced or it is not. However, the processes involved in a phonemic production can be individually rated on a multi-point scale.

Though perceived competence is without degree, an electronically measured competence factor can show degrees of degradation or improvement. Such devices can also rate acceptable and unacceptable phonemic production. An example of such a device is the Alto SIGMATOMETER described in Appendix B. This device measures competence of the surd fricative (s) and its sonnet counterpart (z) by a meter display of the absolute number of axis zero crossing present. A modification of this instrument has been employed for studying one index in determining speech changes under G stress.

Peak power asymmetry offers a second index of speech performance under stress. Since speech represents one of the higher order, non-vital, functions performed by man, it is not unreasonable to assume that disturbing influences will affect speech processes long before many vegetative functions are affected. Asymmetry, being the result of phase structure, is sensitive to those speech functions that would influence phase. Of particular importance is the coupling and size of cavity resonators within the vocal tract. Appendix C describes the Alto VOISCON, an instrument employing peak power asymmetry as a prime method in the distinguishment of voice from noise.

Rate and level of speech are additional parameters which can be employed as separate indices of speech. In addition, level is employed as a normalizing factor in determining absolute asymmetry.

3.3 Factors Affecting Phonation and Fricatives Under Stress

It appears reasonable to assume that emotional disturbances will accompany physical disturbances due to acceleration and to a degree largely dependent on the particular pilot and the situation which he encounters. The effect of these disturbances on speech have been considered as a single combined disturbance. The purpose of this section is to consider briefly several major functional areas that would be disturbed under stress and project possible effects on the production of fricatives and phonation.

a. Respiration: Because a controlled air stream is required for the production of all speech sounds, changes in this function would reflect changes in phonation ability most probably by intensity and/or rate variations. Fricatives (and plosives) require a larger expelled air volume than do phonated sounds so that intensity variations could be expected; however, if the articulators were not affected, it is unlikely that spectral makeup or axis density of the fricatives would be affected.

b. Auditory Acuity: Since speech sounds are the result of aural perception in a closely controlled feedback loop, phonation and fricative production would bear a direct relationship to the speaker's ability to normally perceive these sounds. Even though neuromuscular control be totally unaffected, the originating instructions from the brain would not be correctly applied if the input (aural) information is in error.

c. Neuromuscular Coordination and Control: If the position of the articulators either cannot be adjusted or cannot be adjusted rapidly enough, major changes in asymmetry patterns during phonation, and axis density patterns during fricative formation would be expected. Coordination and contol disfunctions can result due to physical and/or emotional disturbances, and it is believed that tonicity of vocal tract surfaces may be greatly altered during such disturbances which would bear heavily on asymmetry measurements even with normal articulation.

3.4 Objective Speech Measurements

Historically, many investigators have been seeking methods that would separate and/or classify communicative sounds emanating from the human vocal system. A large percentage of such effort has been directed along the line of spectral analysis of speech signals, wherein a measure of energy per unit bandwidth can be obtained over the frequency range of the human auditory tract.

From this and other work, a basis for understanding, and even mathematical analysis, of the processes whereby sounds are produced, and perceived has resulted. Further, machines have been constructed which, within limitations, can generate and decode speech sounds.

Because of the apparent wealth of information contained in the spectra of speech, some recent investigators have employed the "self-organizing system" in an attempt to classify sounds, or more accurately, to classify the spectral patterns which can be derived from sounds. Such machines are "trained", in that the correct answer is supplied along with the information pattern, e.g., the spoken word "five" would be accompanied by manually depressing a button labeled "five". When a second pattern is presented along with the proper training, the machine tends to classify by weighting heavily the unique information in the second pattern and weighting lightly or not at all that information common to both patterns. Is this process similar to the human counterpart in hearing, learning and understanding? Certainly, spectral information is a very important aural input, however, the learning the understanding mechanisms of the human brain and its associated neural transducers are only vaguely known at best; the point being, that even if machines could duplicate "nature's" way of accomplishing an end result, this may not at all represent the best or most efficient solution. The classical example is the aeroplane which, unlike the bird, does not flap its wings.

As an outgrowth of work performed at IBM, Alto Scientific Company has undertaken a series of investigations of speech sounds in manner not necessarily performed by the human hearing mechanism. One important outcome of these investigations is the measurement of a characteristic of speech sound waves termed "peak power asymmetry". The principle involved is one of phase structure within a complex waveform (a characteristic of sound not discernible by the human auditory makeup) and can be measured by the algebraic subtraction in real time of the positive going peak envelope from the negative going peak envelope of an input speech (or other) waveform. A second outcome of Alto research has been the use of axis density measurements in classifying speech sounds. This principle involves the measurement of the number of times an input waveform crosses a reference level (normally the zero reference level) in a given time interval.

The meaning of such parameters as peak power asymmetry and axis density as they relate to the solution of a problem can, as yet, only be empirically determined. Devices employing these principles are already being marketed in several areas: Language Teaching Aids, Therapeutic Applications Involving the Correction of Speech Deficiencies, Devices Responding to Limited Vocabulary, Devices that Recognize the Human Vocal Fold Vibration in the Presence of High Ambient Noise Environments, and Speech Teaching of the Deaf.

These principles have been implemented in terms of hardware and used in performing this study at the Ames Five-Degree-of-Freedom Centrifuge in an attempt to discern changes of voice characteristics as a function of G stress.

3.5 The Heart and Heart Sounds

It is generally agreed that the easily detectable cardiac sounds are a result of muscle contraction and valving. The energy contained within the high frequency region (above 500 cps) that represent minor heart sounds are considered clinically important in certain cardiac disfunctions such as murmurs. However, since little contribution to peak power asymmetry measurements results from such sounds, only the lower frequency, higher energy sounds will be considered.

The heart is basically a compartmented muscle possessing four special properties:

a) It is automatic not requiring stimuli from the central nervous system for maintenance of its activity (although the nervous system may modify its activity).

b) It is rhythmic displaying an inflexible series of operations or events.

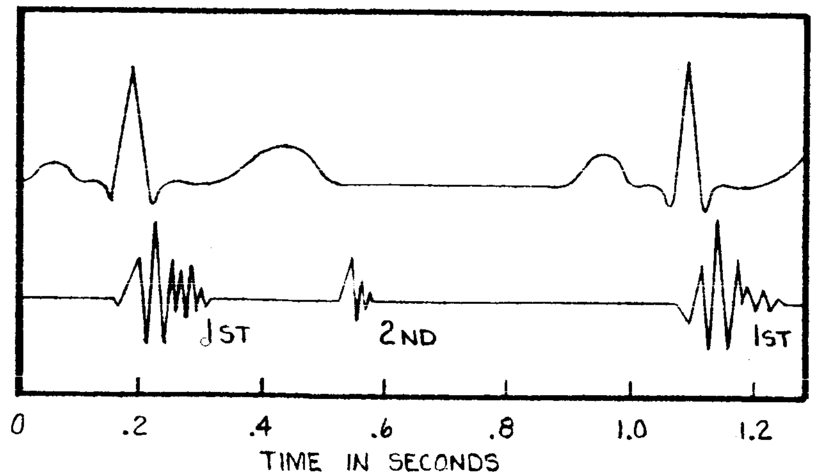
c) It is insensitive during its contraction cycle to further stimulus so that its rhythmicity is not disturbed.

d) It is binary in response either making a maximal contraction or none at all.

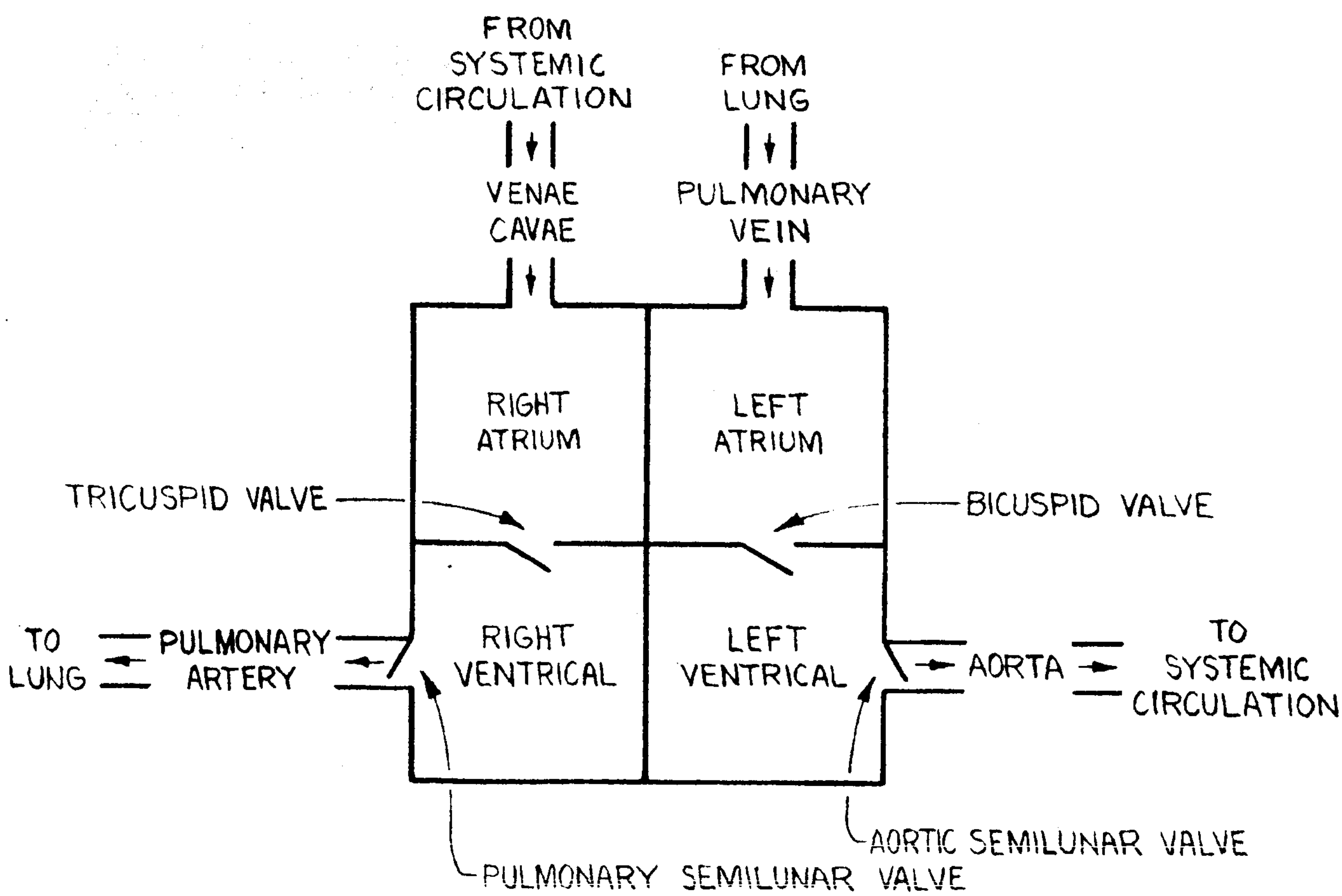
Referring to Illustration 2B, a vastly simplified description of a cardiac cycle follows: Ventricular contraction causes closure of the tricuspid and bicuspid valves. Blood is thereby pumped into the pulmonary artery to the lungs through the pulmonary semilunar valve, as well as into the aorta to systemic (or greater) circulation through the aortic semilunar valve.

Oxygenated blood from the lungs via the pulmonary vein; as well as venous blood from systemic circulation, empties into the left and right ventricles respectively, where, during atria contraction, blood is pumped into each respective ventricle.

Generally, two cardiac sounds are easily recognized: The first sound as a result of ventricular contraction and bicuspid/tricuspid valve closure, and the second sound as a result of semilunar valve closure at the time of ventricular relaxation when arterial blood pressure exceeds ventricular pressure. Illustration 2A (bottom trace) depicts typical recorded cardiac sounds. The first sound is of lower pitch and longer duration than the second sound; for



A. TYPICAL EKG AND PHONOCARDIOGRAM.



B. SCHEMATIC OF THE HEART

ILLUSTRATION 2

this reason, the second sound is normally acoustically perceived as the loudest.

The top trace of Illustration 2A depicts an EKG signal and its time relationship to the first and second cardiac sounds.

3.6 Factors Affecting Heart Sounds Under Stress

Empirical determinations made in the past indicate that most impact type sounds display peak power asymmetry. Cardiac valving, which contributes to both the first and second cardiac sounds, tends to generate a hydraulic impact due to the sudden interruption of flow. As can be seen in Illustration 2A (bottom trace) a definite positive valve of peak power asymmetry is present. If, under G stress, flow interruption differs from that at 1 G, it would be expected that the absolute magnitude of asymmetry may also differ. Further, if the hydraulic damping caused by the elasticity of the arterial system is in any way influenced by G stress (or by physical or emotional disturbances as a result of G stress) absolute asymmetry would also be affected.

4.0 Data Accumulation and Reduction

Installation of measuring equipment was performed at the Ames Five-Degree-of-Freedom Centrifuge between 6-24-63 and 6-27-63. Eighteen pilots in a total of 158 runs were monitored during the period of 6-6-63 and 6-27-63 to 7-19-63. Appendix E presents a chronological history of dates, number of runs monitored, and pilots.

In the accumulation of data, certain constraints were imposed in that the prime purpose of the runs monitored was not the objectives of this study. Briefly, the constraints were as follows:

- o Accumulation of data on a non-interference basis.
- o Mandatory use of a Roanwell microphone which was connected into the intercom system.
- o Unwillingness of some pilots to speak during runs*.
- o Unwillingness of some pilots to wear the cardiac transducer.
- o Majority of data runs made at 3.5 G or less.
- o Inability to optimize placement of cardiac transducer
- o Addition of vibration to G bias.

Initially, it was anticipated that only 20 runs would be monitored, however, because of the constraints, about 8 times this number were monitored in an attempt to find cooperative subjects which would recite test sounds at 1 G and the same test sounds at higher G levels (without vibration). In addition, in order to obtain cardiac information with as few extraneous influences as practical, subjects under G stress should remain physically and vocally quiet.

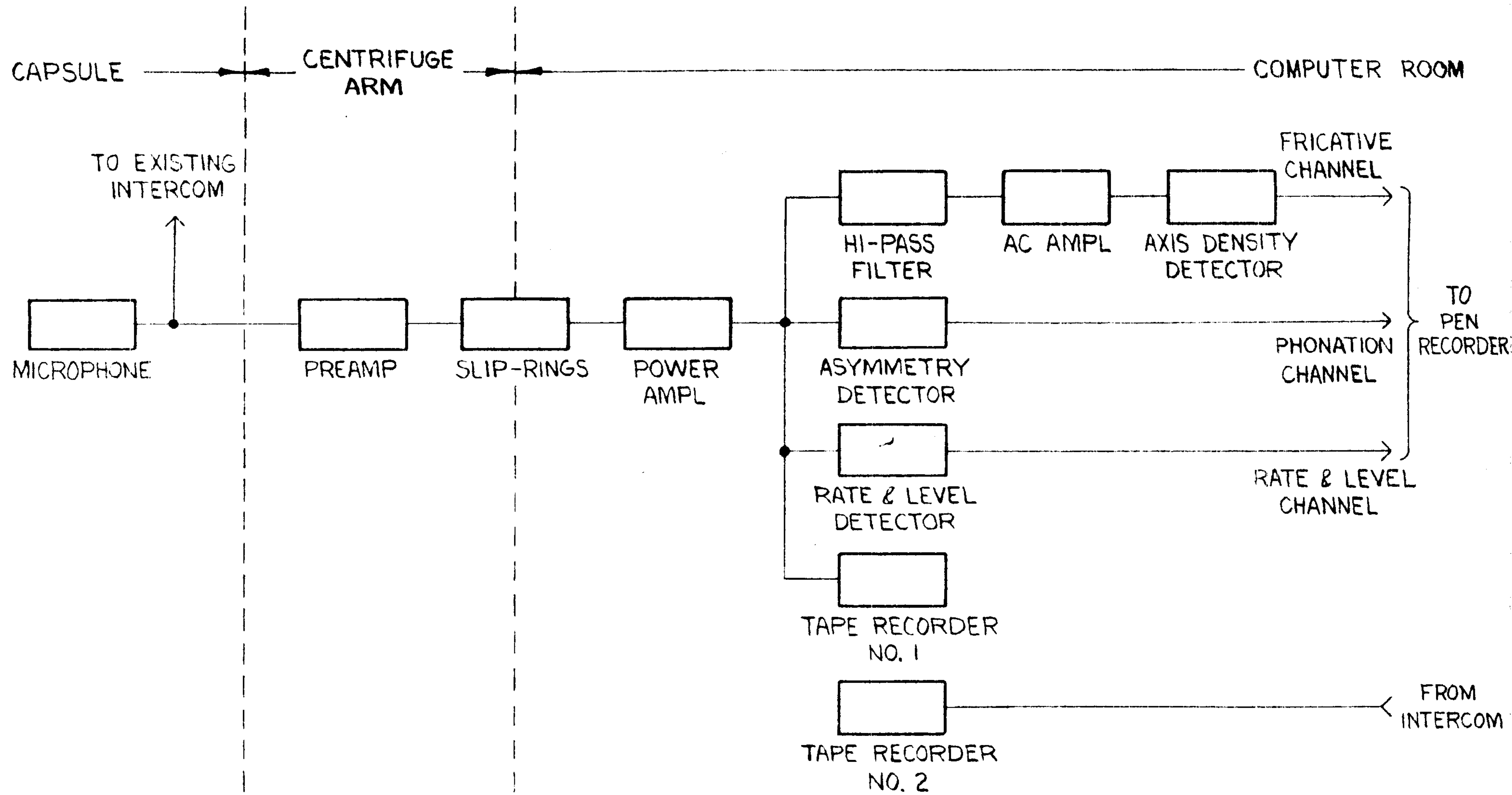
Parenthetically, no criticism of any pilot is implied or intended. Alto is particularly grateful to Mr. R. Weick and the centrifuge crew all of whom cooperated to the greatest extent possible.

4.1 Measuring Equipment, Voice Channel

Referring to the System Block Diagram, Voice Channel, Illustration 3, the following equipment was recommended for use:

- o Microphone - Electrovoice 624
- o Preamplifier - Shure M65 (modified)
- o Power Amplifier - Scott 299C
- o Axis Density, Asymmetry and Rate/Level Detectors - Alto
- o Tape Recorder #1 - Ampex 601
- o Tape Recorder #2 - Wollensac T1515-4

*Due to possible distraction in performing other problem assignments.



SYSTEM BLOCK DIAGRAM, VOICE CHANNEL
ILLUSTRATION 3

Of the equipment recommended, the microphone was the only item changed by Ames because of unavoidable circumstances. A Roanwell M87/A1C microphone was substituted, which necessitated the addition of a hi-pass filter (Allison 2A) and an AC Amplifier (HP 450) in the fricative channel. The substitution of the Roanwell microphone was unfortunate for the following reasons: a) Asymmetry output of this microphone was essentially unidirectional and of very small magnitude. Such characteristics are the result of poor low frequency response, a prestressed diaphragm, and high Q resonances in the low frequency portion of the pass band. b) Fricative axis density measurements are best made above 7 KC. The high frequency cut-off point of the Roanwell microphone was approximately 3 KC. This necessitated the addition of a filter/amplifier in order to attempt to separate the higher order phonation formants from the relatively low energy content of the fricatives and still remain within the passband of the microphone

Because of the impractibility of utilizing the Electrovoice microphone, it was necessary that Alto equipment be tied in parallel with the input to the centrifuge intercom amplifier. Noise and hum from this tie-in prevented attainment of signal-to-noise ratios that would have permitted high quality recordings.

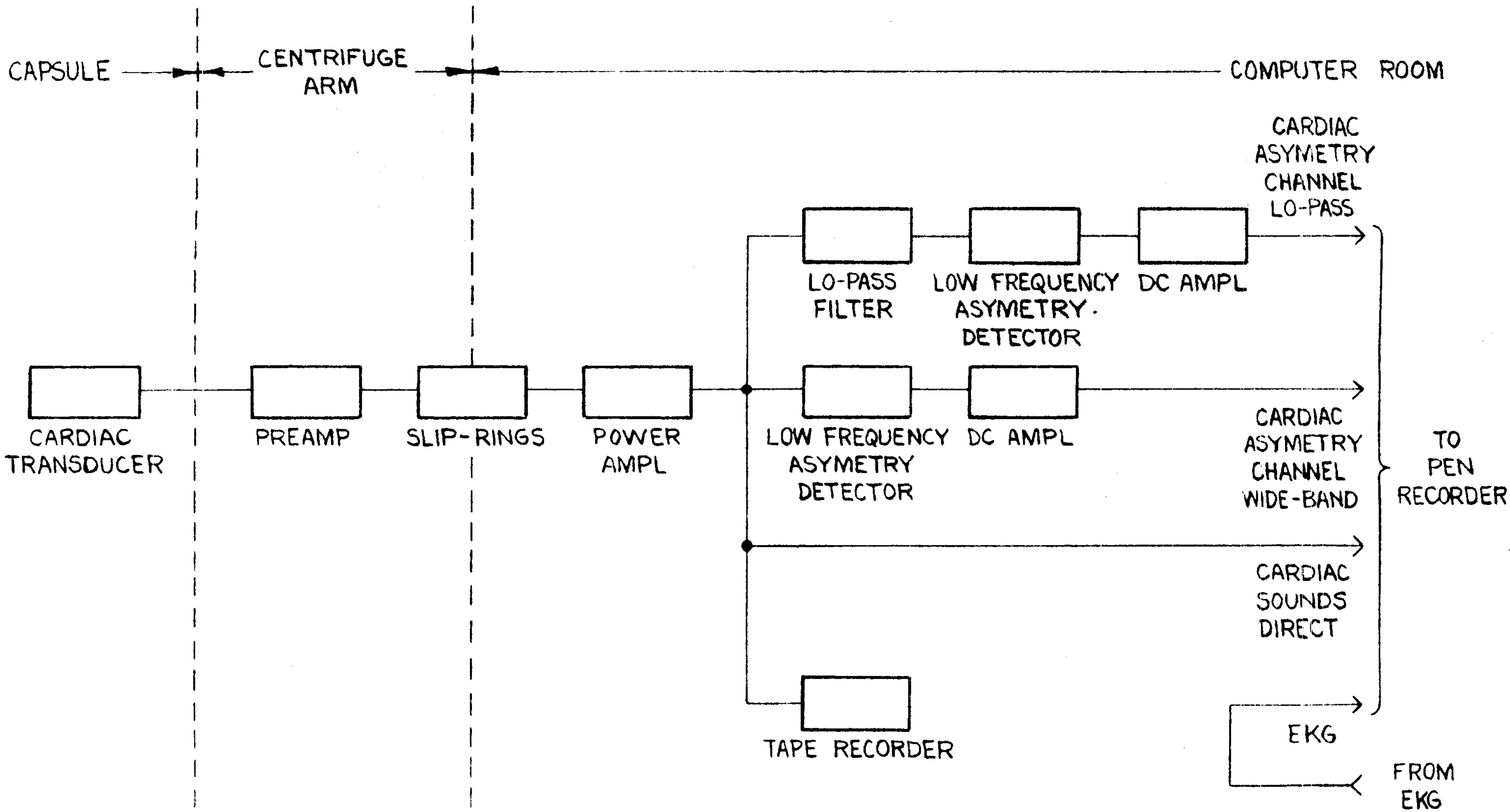
4.2 Measuring Equipment, Cardiac Channel

Referring to the System Block Diagram Cardiac Channel, Illustration 4, the following equipment was recommended for use:

- Cardiac Transducer - Clennite MP202
- Preamplifier - Shure M 65 (modified)
- Power Amplifier - Scott 299C
- Low-Pass Filter - Alto 150 cps linear phase LP Filter
- Low Frequency Asymmetry Detector - Alto
- DC Amplifiers - Dymec 2460A and Kane KEL-506
- Tape Recorder - Ampex 601

Of the equipment recommended, the cardiac transducer was the only item changed because of excess cable noise in the capsule wiring. A lower impedance transducer was substituted (Sanborn 350-1700-C10) in order to obtain data. The major disadvantage of the Sanborn is the relatively large degree of acoustical coupling of ambient background sounds via air media. In addition, since the Voice-Cardiac System employed a common ground, noise and hum were also introduced into the cardiac channel for the reasons listed in Measuring Equipment, Voice Channel.

A photograph of the electronic equipment is shown in Illustration 5.



SYSTEM BLOCK DIAGRAM, CARDIAC CHANNEL
ILLUSTRATION 4

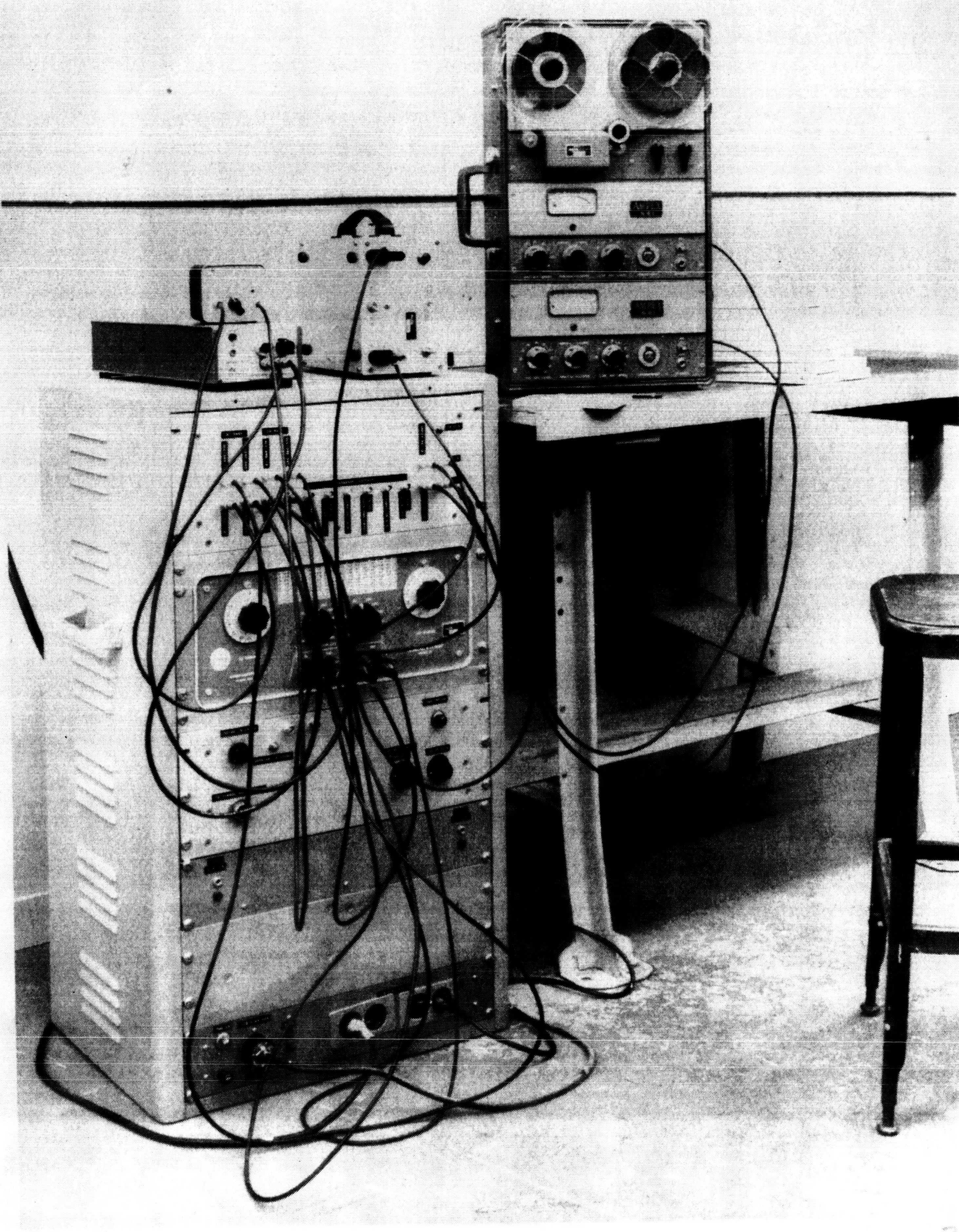


ILLUSTRATION 5

4.3 Speech Data

Upon completion of the test runs, a chart was recorded at Ames (by Alto personnel) using an Electrovoice 664 microphone in place of the Roanwell microphone in order to illustrate the type of baseline data originally expected. (See Illustration 6A.) Note that the asymmetry of those vowels representing small oral cavity openings, (i) and (u), tend to display opposite polarity from those representing the larger cavity openings. In Illustration 6B, a number count of "one" through "zero" is shown. The fricative channel shows deflections for both the initial and final (s) in the word "six" and the initial (s) and (z) sounds in the words "seven" and "zero" respectively. Further, the fricatives do not show deflections on the asymmetry or rate/level channels since they are symmetrical and of considerably lower energy level than the vowels.

Referring to Illustration 7, typical sections of the better chart recordings are shown. In Illustration 7A, data at rest was taken from McMillan (Data run #9C) in the form of a number count of "one" to "seven" to "one". Asymmetry deflections are primarily unidirectional, whereas fricative and rate/level channels appear as expected. The same number count was repeated at 3.5 G bias with vibration interposed during half the count (Illustration 7B). Substantial changes in data appear in two ways: first, the average level with which the words are produced approximately doubled; second, the fricative channel responded to each word. A rather large noise content on the asymmetry channel prevented any quantitative reduction of data; however, the approximate asymmetry, when nomalized for level variations, does not show any discernible trend at this level of G bias. Because vocal pitch generally increases with level, the output of the fricative channel under G bias is most probably due to increased energy level of the higher order vowel formants. The use of the limited bandwidth Roanwell microphone is predominately responsible for permitting high order vowel formants to appear as fricatives, although the power amplifier dynamic range limitation may be approached during some words which would result in increased distortion.

Illustration 8 depicts data taken from Gary McNabb (Data run #16D) at 3.5 G bias with vibration interposed during the number count. Baseline data was not available from this pilot; however, the similarities to Illustration 6B are apparent.

Of all voice data accumulated, one particular run was made at 6 G without vibration. In addition, this was the only run observed in which the controlled vocabulary of sounds was spoken at 1 G and again at 6 G. Unfortunately, not all electronic equipment was installed at the time of this run so that data was tape recorded and the results photographed from an oscilloscope. Referring to Illustration 9, the asymmetry and rate/level is shown for each of the two G conditions. The traces are interpreted as

SUBJECT: THOMAS MOORE

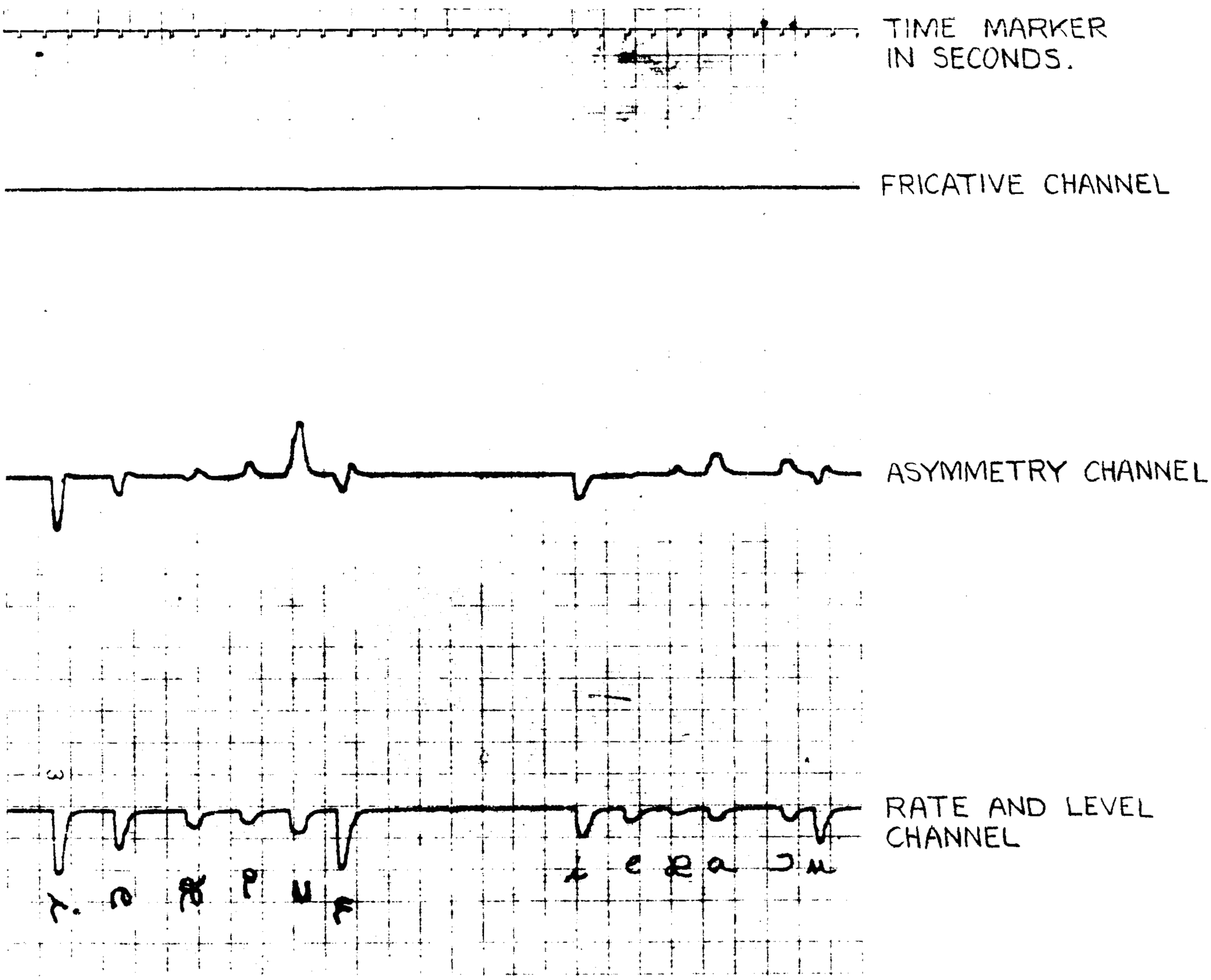


ILLUSTRATION 6A

SUBJECT: THOMAS MOORE

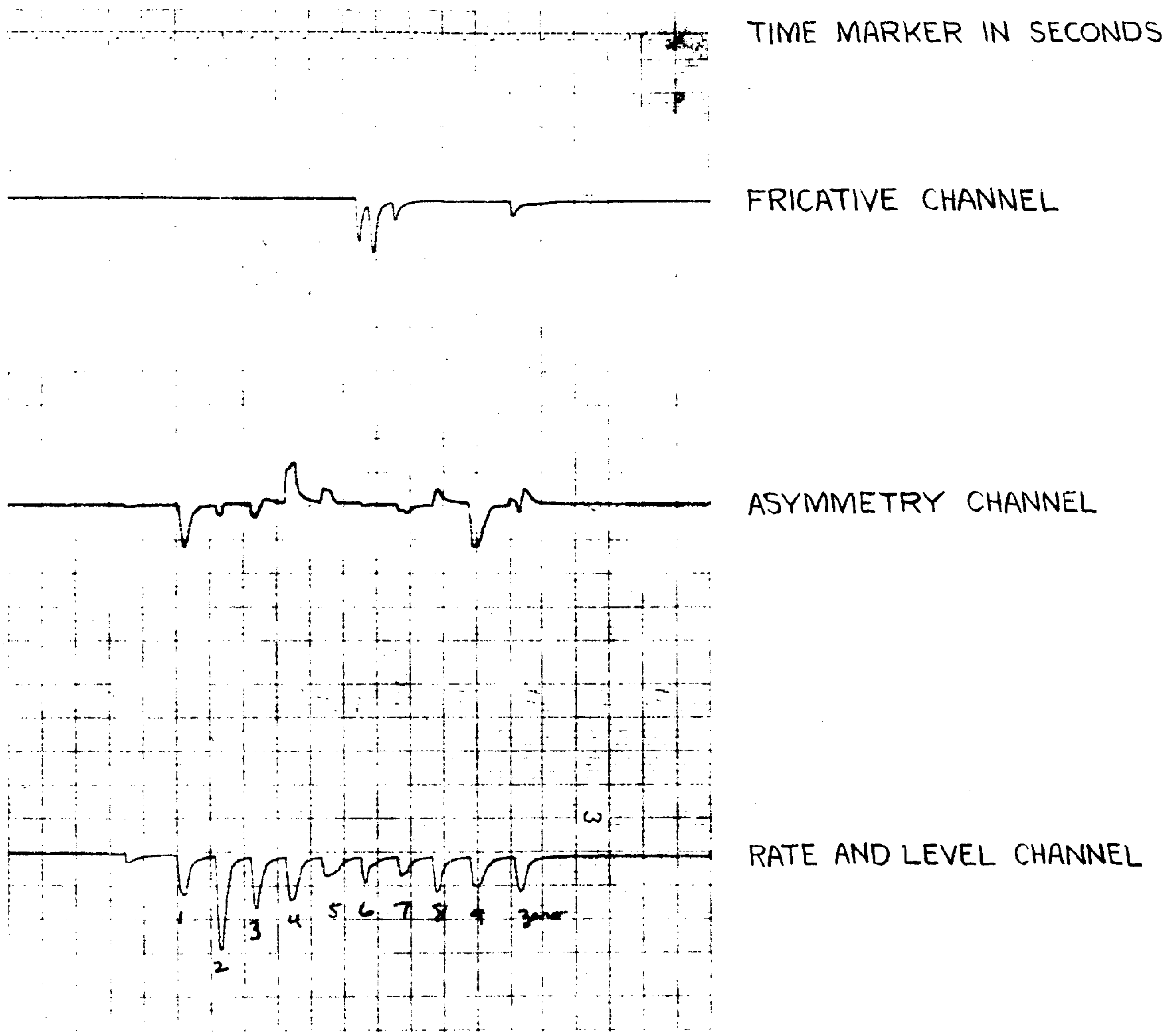


ILLUSTRATION 6B

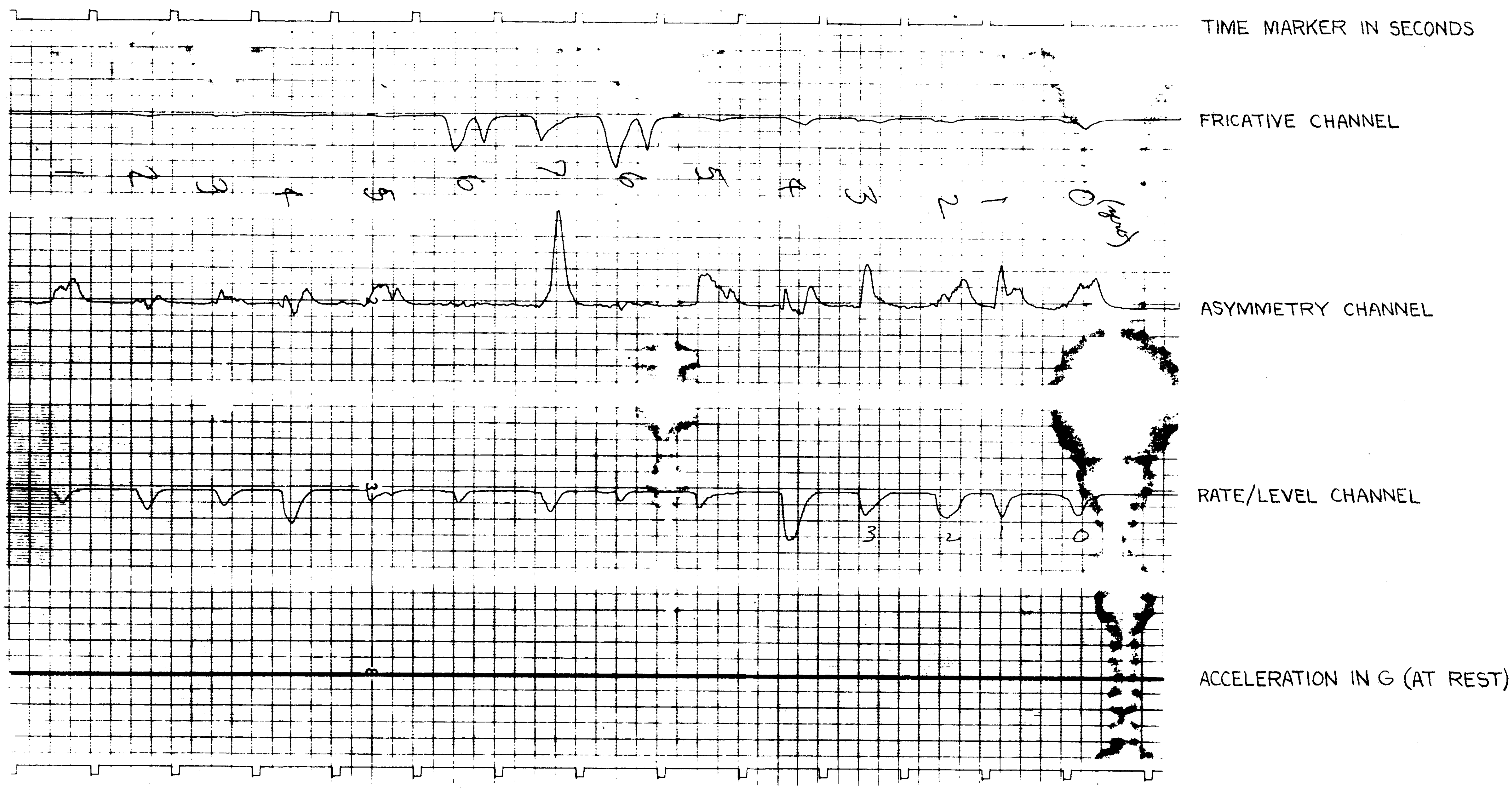


ILLUSTRATION 7A

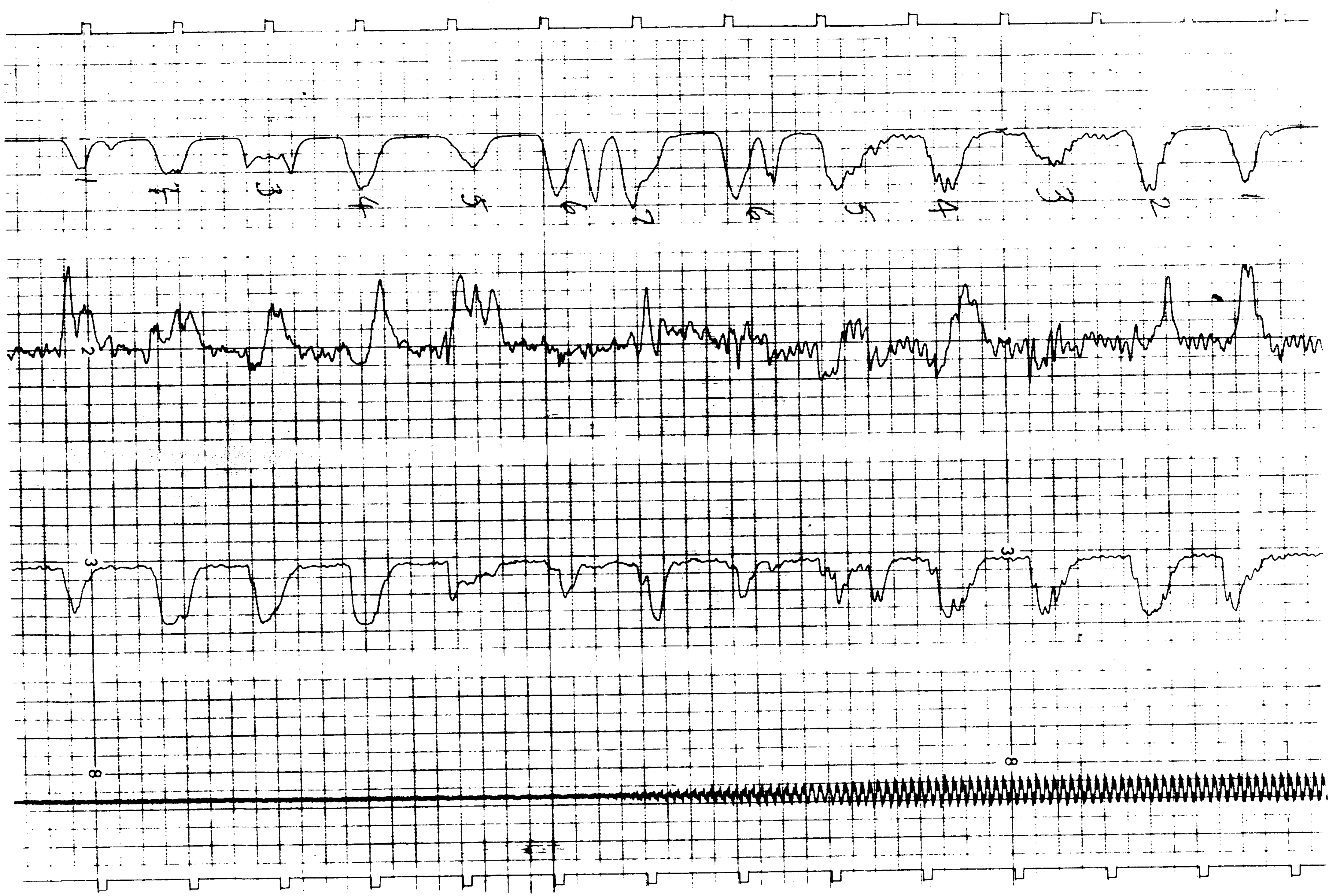


ILLUSTRATION 7B

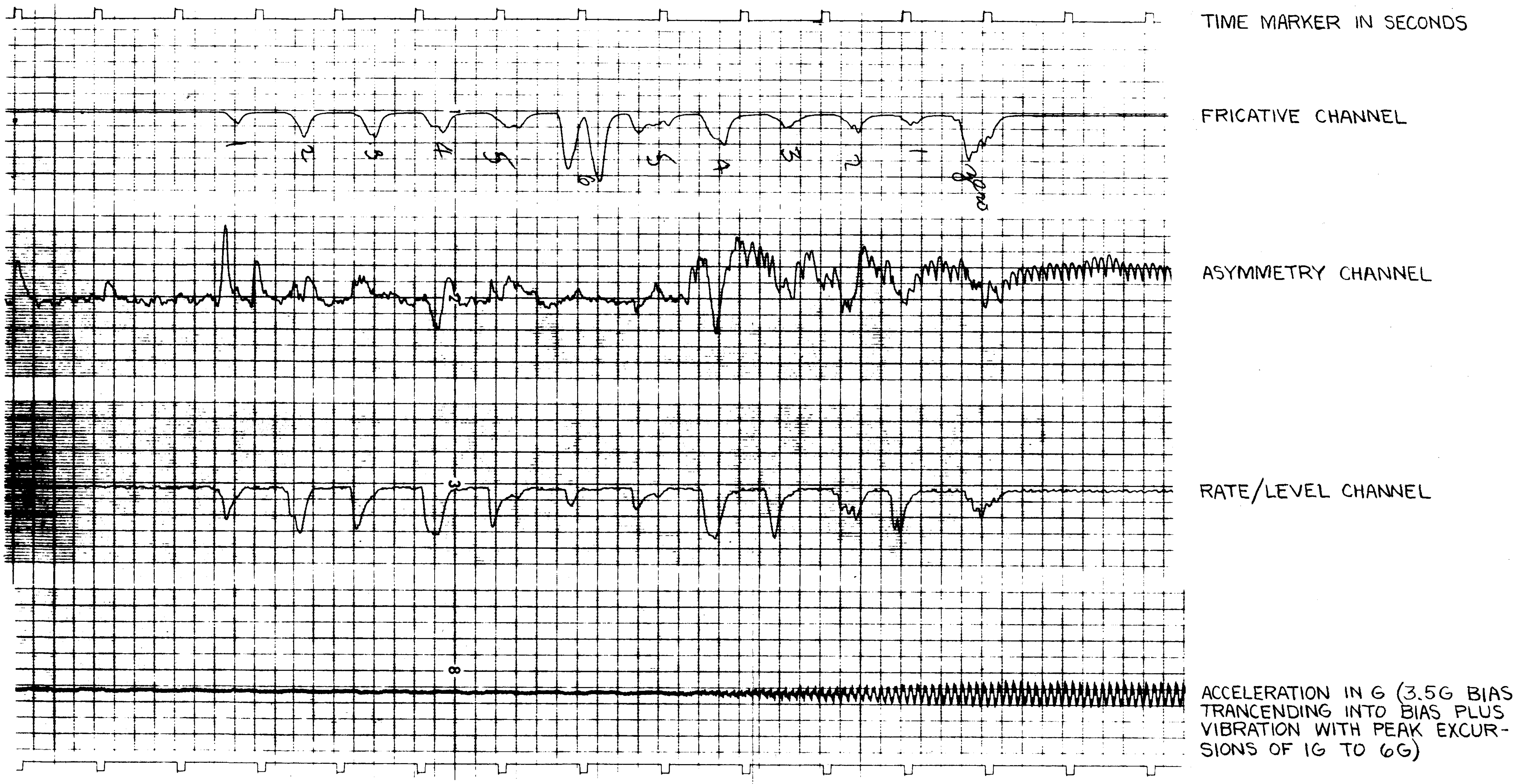
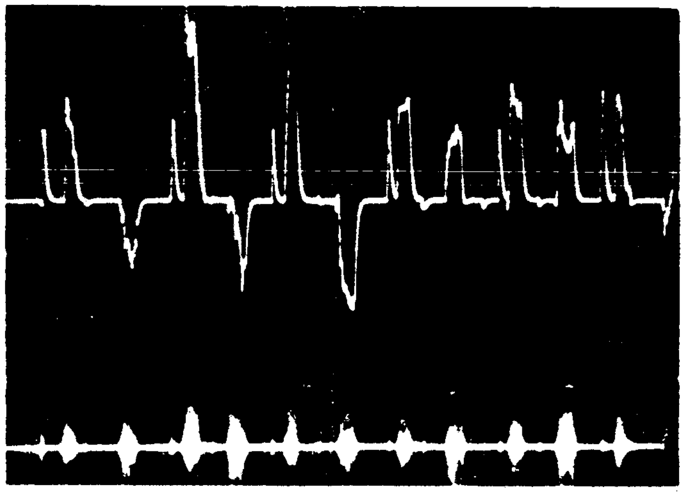


ILLUSTRATION 8

6G RUN AT AMES 6-6-63
PILOT: GLENN
REPRODUCED FROM TAPE

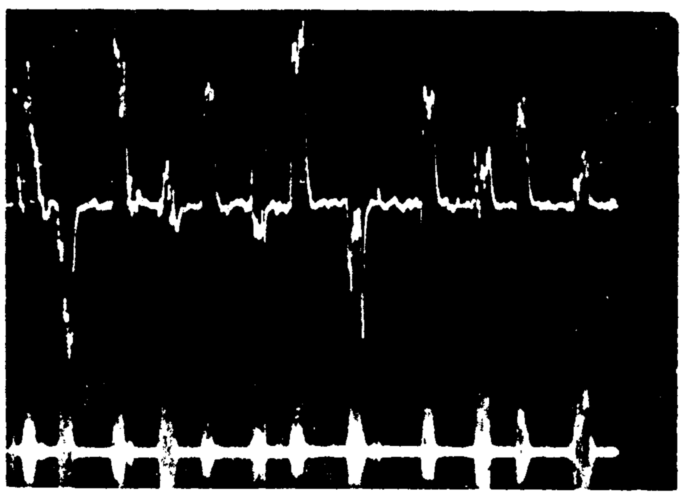
(i) | (e) | (æ) | (a) | (ɔ) | (u)



T T T T T T

DATA AT 1G

(i) | (e) | (æ) | (a) | (ɔ) | (u)



T T T T T T

DATA AT 6G

ILLUSTRATION 9

follows: Mr. Thomas Moore of Alto first phonated each sound; after a short pause the sound was repeated by the pilot. The time intervals labeled "T" are the reference sounds (in some cases preceded by the word "now" or "and"), whereas the unlabeled intervals are the pilot produced sounds. Because subjective data changes are discernible between the two G conditions, this data was reduced by deriving an index number representing the asymmetry for each sound. Since the absolute magnitude of asymmetry, when present, is a first order function of level, an index number will result from the following operation:

$$I_{(\underline{\ \ })} \qquad \frac{\int_{t_1}^{t_2} A_{n}\, dt}{\int_{t_1}^{t_2} L_{0}\, dt}$$

where $I_{(\underline{\ \ })}$ is the index for a particular phonetic sound

A is the asymmetry envelope, the subscript being the G bias

L is the level envelope, the subscript being the G bias

$t_2 - t_1$ is the length of phonetic production

Results are as follows:

		1 G	6 G
$I_{(\underline{i})}$	=	-.53	-.92
$I_{(\underline{e})}$	=	-.062	+.012
$I_{(\underline{æ})}$	=	-.62	-.013
$I_{(\underline{a})}$	=	+.36	-.22
$I_{(\underline{ɔ})}$	=	+.76	+.35
$I_{(\underline{u})}$	=	+.21	+.24

Because of the limited nature of this data, other methods of reduction and normalization have not been attempted. In addition, each sound spoken represents a phoneme so that additional data is necessary in order to determine the separate sensitivity of the index number due to phonemic production and stress.

4.4 Cardiac Data

Cardiac data was obtained from most pilots during the time voice data was in the process of being taken. A typical phonocardiogram was made at Alto using the same equipment ultimately installed at Ames. Referring to Illustration 10, a phonocardiogram typical of those to be expected along with the wide band and lo-pass asymmetry patterns are shown. Asymmetry is definitely evident though the physiological reason for its presence is beyond the scope of the present study. The top trace displays unprocessed heart sounds; the second trace displays asymmetry of those frequency components up to 150 cps; the bottom trace displays wide-band asymmetry. (Note that the inversion of the bottom trace with respect to the second trace is a function only of an inversion by the oscilloscope photographed.) No attempt was made to interpret such data in itself; instead, comparison of such data at different levels of G stress was contemplated in order to discover if cardiac asymmetry was in any way related to stress.

Illustration 11 and 12 depict typical sections of cardiac chart recordings made, Illustration 11 taken from pilot Jim Show (Data Run 019C) depicts cardiac data at rest. Illustration 12 depicts cardiac data from the same pilot during the application of G bias. The addition of vibration, not shown in these runs, simply increased the noise level present.

Due to system hum and noise as well as superimposed vibration during G bias, preliminary examination of the cardiac data recorded (to 3.5 G) offered very little promise in the subjective determination of any changes. In exploring useful methods of data reduction, several comparison photographs were taken in the Alto laboratory. From these photos, one unexpected result has been partially uncovered, i.e., under certain circumstances with the subject tested, the absolute polarity of measured asymmetry appears to reverse as the respiratory cycles transcends from inhalation to exhalation.

This data is presented on only a tentative basis, for the scope of the present contract does not provide for additional investigation in this area; further, effects of increased cardiac-monitoring point distance during inspiration of breath (which would influence amount but not necessarily polarity of asymmetry) and increased transducer diaphragm pressure during inspiration have not been analyzed. Referring to Illustration 13, definite respiratory patterns are shown where the subject breathed deeply at a controlled rate.

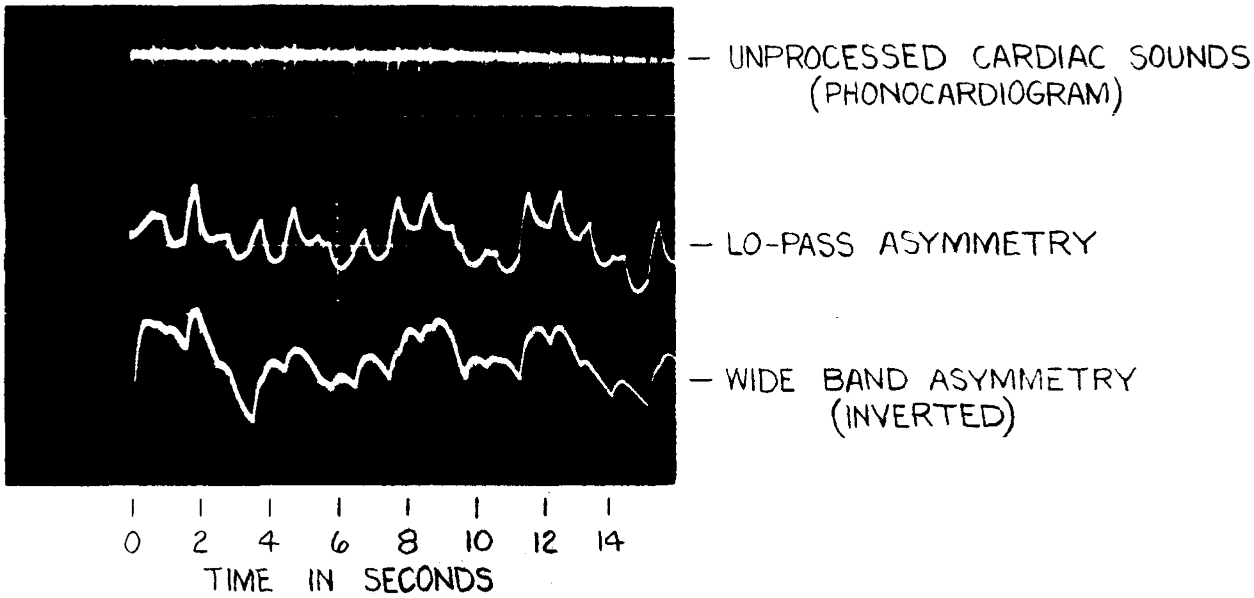


ILLUSTRATION 10

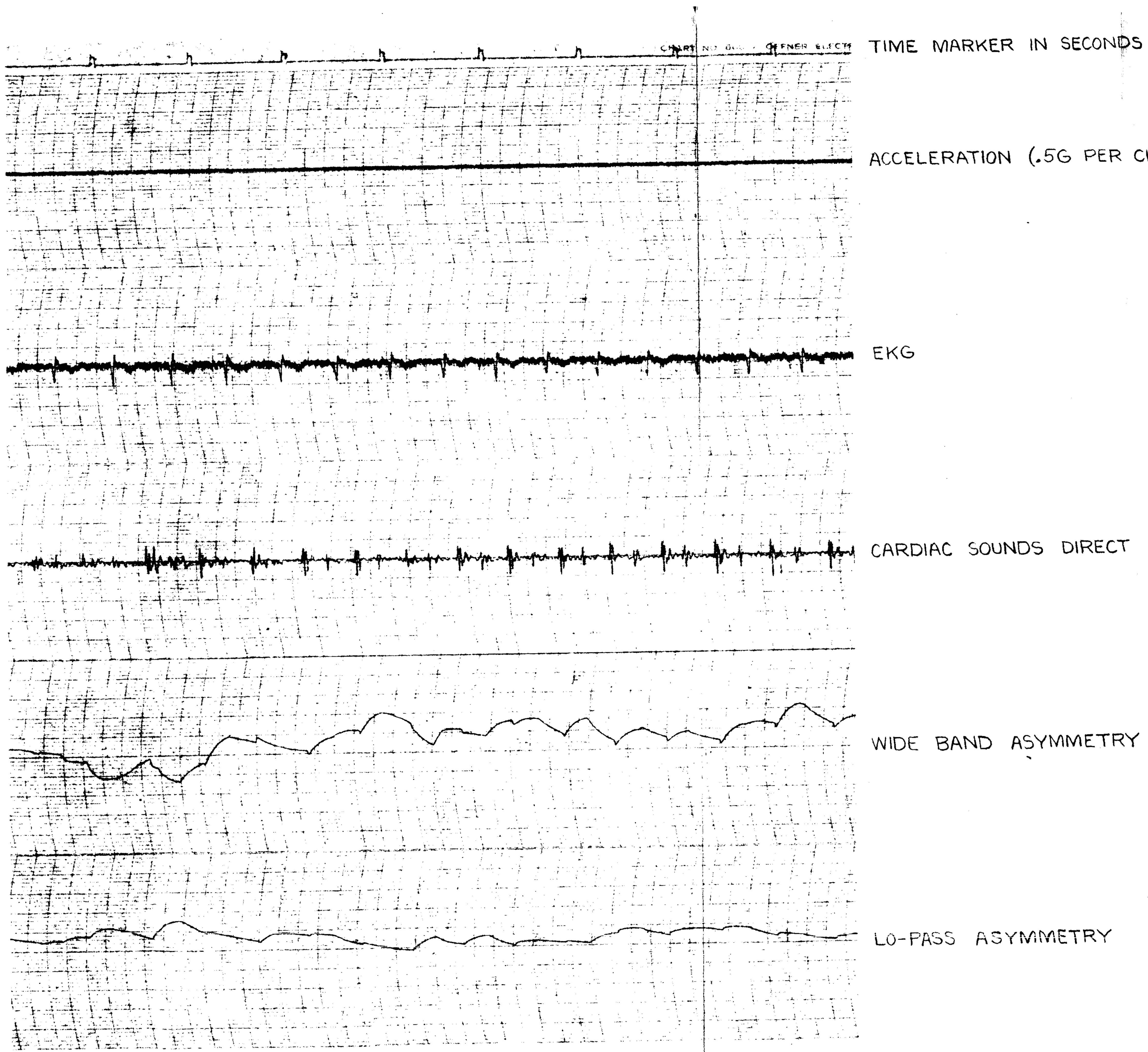


ILLUSTRATION II

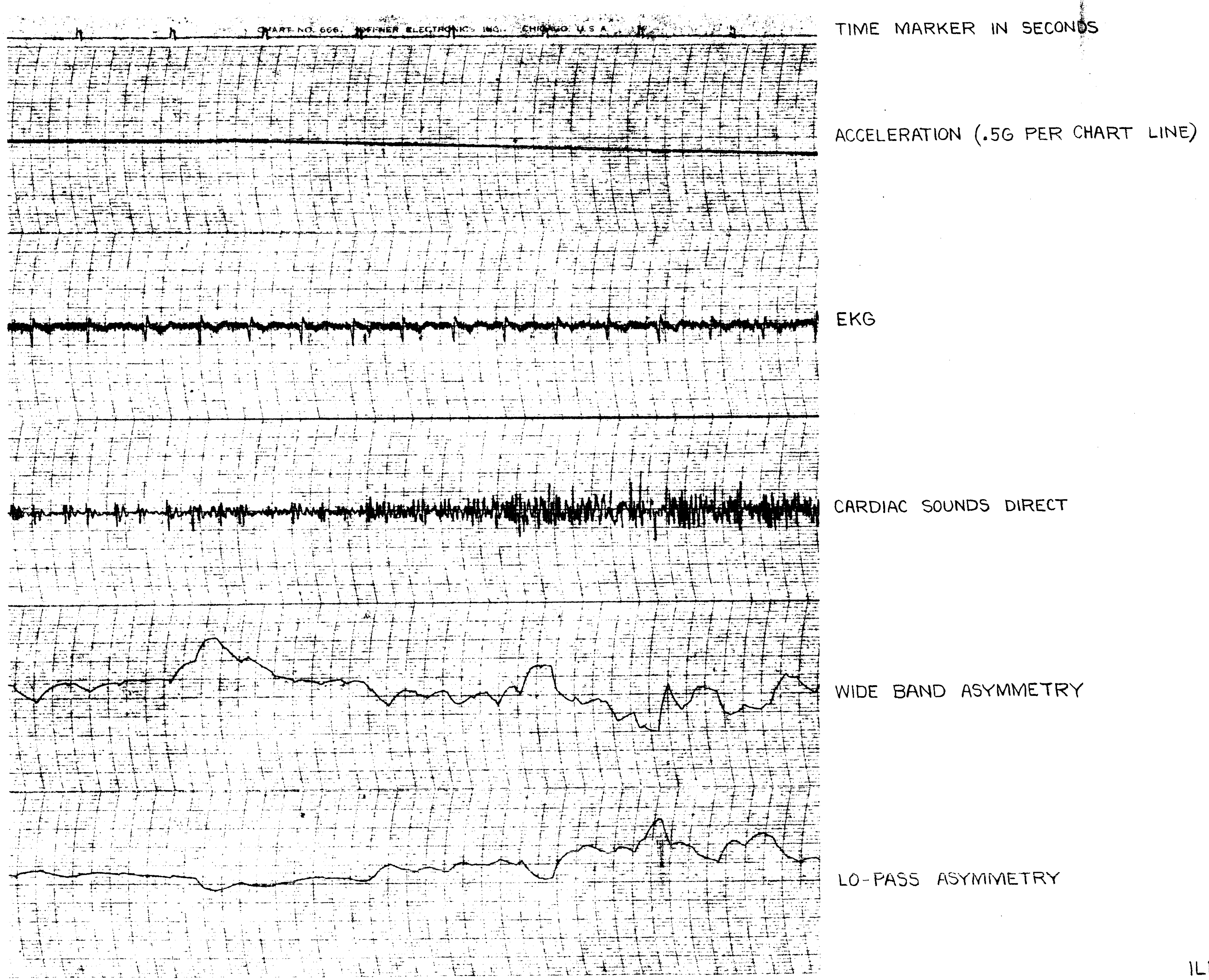


ILLUSTRATION 12

5.0 Conclusions and Recommendations

5.1 Realization of Objectives

- o Speech sounds - Of the large amount of data taken, that most suitable for preliminary reduction was at the higher (6 G) level. If trends, other than a level increase, are present at the lower (3.5 G) level, they have been masked by the microphone dictated for use and the lack of control which Alto was able to excise over experimentation (due to the "piggy-back" nature of this study). The data taken at 6 G displays marked changes in asymmetry from that at 1 G. As a result, a tentative conclusion might be drawn that voice sounds, when viewed on the basis of asymmetry, change little or not at all up to some given level of G bias (presumably between 3.5 G and 6 G) and change rapidly after some threshold point.

- o Cardiac Sounds - Because of constraints in the accumulation of data, no attempt has been made to reduce cardiac data from actual centrifuge runs beyond the conclusion that the heart sounds do display asymmetry. Limited laboratory data may indicate a relationship between cardiac asymmetry and respiration.

5.2 Recommendations

In order to fully investigate the possibility of rating "pilot well-being" by means of his speech, and to obtain more information concerning the relationship of cardiac asymmetry to stress, it is recommended that a series of centrifuge experiments be conducted solely for this purpose. In addition, controlled experimentation at incremental G levels up to as high as 12 or 14 G with the inclusion of physiological and bio-medical monitoring to the greatest feasible extent should allow data correlation not possible under the current effort.

Certain other areas might be worthy of simultaneous investigation during controlled conditions, e.g., the feasibility of man-machine communication via speech, and the measurement of other speech and cardiac parameters as further indices of well-being (spectral distribution and frequency derivatives).

It is recommended that any follow-on study be separated into two phases as follows:

Phase I

- o Construct measurement equipment in "building block" form so that circuit optimization is possible in the centrifuge environment.
- o Make the following measurements as a function of G stress by modification of Alto speech responsive devices:

    - Real time rate and level as an analog function of time.

SUBJECT: THOMAS MOORE

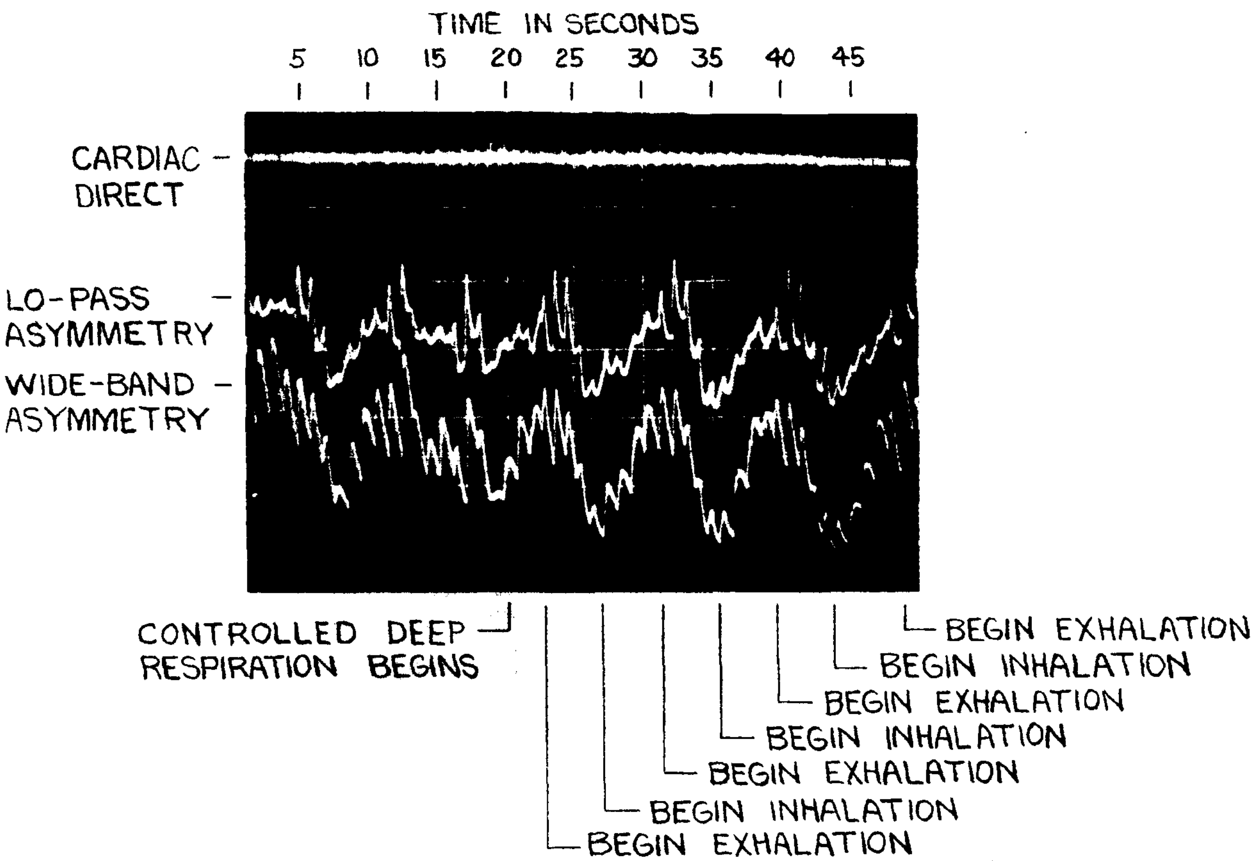


ILLUSTRATION 13

- Real time fricative and selected plosive competence as an analog function of time.
- Asymmetry integrated over the test phrase.
- Fricative and selected plosive integrated over the test phrase.
- Rate and level integrated over the test phrase.

o Make the following cardiac measurements as a function of G stress by modification of Alto speech responsive devices:

- Real time phonocardiogram.
- Real time cardiac asymmetry as an analog function of time with consideration of predetection filtering to obtain best results.
- Cardiac asymmetry integrated over a specified number of EKG samples.

o Investigate certain other speech parameters as a function of G stress on a preliminary basis.

- Change in spectral density of fricatives and affricates.
- Change in tonal quality of phonation.

o Include certain voice circuits as ancillary equipment in order to obtain preliminary data concerning the feasibility of direct control via a speech function.

Phase II

o Design and construct an adaptive threshold logic computer to operate in conjunction with the speech responsive devices of Phase I.
o Construct equipment for measurement of spectral distribution and frequency derivatives as additional inputs to the computer.
o Make speech/cardiac measurements as a function of G stress and determine by relative computer weighting those parameters that exhibit the greatest relationship to G change.
o Refine those measurement devices exhibiting the greatest G relationship and employ them as sole computer inputs.
o Calibrate computer weighting in terms of G stress so that G level is presented directly as a result of speech input.

APPENDIX A

Table of Phonetic Symbols

APPENDIX A

TABLE OF PHONETIC SYMBOLS

The Neutral Vowel

(ʌ)	up	(ʌp)	

The Closed Vowels

(i)	eat	(it)	
(ɪ)	it	(ɪt)	
(e)	debris	(debri)	
(ɛ)	pet	(pɛt)	
(æ)	pat	(pæ t)	
(a)	ask	(ask)	as often pronounced in America between (æ) and (ɑ)

The Intermediate Vowels

(ɝ)	early	(ɝlɪ)	General American pronunciation
(ɜ)	early	(ɜlɪ)	Southern and Eastern pronunciation
(ʌ)	(See Neutral Vowel above)		
(ɑ)	palm	(pɑm)	

The Open Vowels

(u)	food	(fud)	
(ʊ)	foot	(fʊt)	
(o)	notation	(noteɪʃə n)	
(ɔ)	all	(ɔl)	
(ɒ)	sorry	(sɒrɪ)	as pronounced in England and often in America. Between (ɑ) and (ɔ)

The Schwa Vowel

(ə)	above, below, constitute	(əbʌv), (bəlou), (konstətɪut)

The Vowel Glides (Semi-vowels and Diphthongs)

(w)	witch	(wɪtʃ)
(hw)	which	(hwɪtʃ)
(j)	you	(ju)
(hj)	hue	(hju)
(l)	law	(lɔ)

The Vowel Glides (Continued)

(r)	raw	(rɔ)
(aʊ)	now	(naʊ)
(oʊ)	no	(noʊ)
(eɪ)	day	(deɪ)
(ɔɪ)	boy	(bɔɪ)
(aɪ)	eye	(aɪ)

Laryngeal Modifications Of The Vowels

(h)	hop, hue, which, how	(hap), (hju), (hwɪtʃ), (hau)
(ʔ)	up not down	(ʔʌp nat daun) as sometimes spoken

Oral Modifications Of The Vowels (The Consonants)

The Stop Plosives

(p)	pat	(pæt)
(b)	bat	(bæt)
(t)	two	(tu)
(d)	do	(du)
(k)	class	(klæs)
(g)	glass	(glæs)

The Continuant Fricatives

(f)	fife	(faɪf)
(v)	five	(faɪv)
(θ)	bath	(bæθ)
(ð)	bathe	(beɪð)
(s)	sue	(su)
(z)	zoo	(zu)
(ʃ)	mission	(mɪʃən)
(ʒ)	vision	(vɪʒən)

The Affricates *

(ts), (st)	cats, cast	(kæts), (kæst)
(dz), (zd)	fads, raised	(fædz), (reɪzd)
(tʃ), (ʃt)	church, rushed	(tʃɜtʃ), (rʌʃt)
(dʒ), (ʒd)	judge, rouged	(dʒʌdʒ), (ruʒd)

Nasal Modifications Of The Vowels

(m)	mow	(moʊ)
(n)	no	(noʊ)
(ŋ)	sing	(sɪŋ)

*There are other affricate combinations. These are typical examples.

APPENDIX B

Description of Sigmatometer

for Measurement of Fricative Competence

APPENDIX D

Definition of Terms

Plosive - a term used in the description of a speech sound wherein the release of a pressure buildup occurs.

Pulmonary - pertaining to the lung

Semilunar - shaped like a half moon

Sigmatometer - trade name of a device to measure competence of the production of strong fricative sounds

Sonnet - with voice, a voiced sound

Surd - without voice, a voiceless sound

Systemic - greater or general

Tricuspid Valve - valve joining right atrium and right ventricle and having three cusps (or joints)

Velum - the soft palate

Venae Cavae - Vein emptying into right atrium and formed by the junction of other veins

Ventricle - small cavity

Voiscon - trade name of a device to recognize phonation in the presence of high ambient noise levels

APPENDIX D

Definition of Terms

Affricate - combination of fricative and plosive, e.g., the (st) in "fast" (fæst)

Aorta = The great artery carrying blood from the left ventricle

Articulation - the interruption of the phonated or non-phonated breath stream

Atrium - the first chamber of the heart

Axis Density - the number of times a waveform intercepts a particular reference level

Bicuspid Valve - valve joining left atrium and left ventricle and having two cusps (or points)

Dipthong - a double vowel in a sterotyped combination

Formants - frequency bands in which energy occurs during production of sound

Fricative - a term used in the description of a speech sound when friction sound waves (random gaussian noise) are a part of the total sound wave complex

Glide - a term used in the description of a transitory sound produced during movement of articulatory structure from one position to a second position

Larynx - anatomical section in which sounds are produced (voice box)

Mandible - the lower jaw

Nasal - a term used in the description of a speech sound having nasal resonance as its peculiar acoustic property

Peak Power Asymmetry - the real time subtraction of the envelope of positive going portion of a waveform from the envelope of the negative going portion

Pharynx - space behind the nose and mouth

Phonation - laryngeal vibration (vocal fold vibration)

Phoneme - a sound family

Pitch - fundamental frequency of phonation

APPENDIX E

Chronology of Centrifuge Data Runs

APPENDIX E

CHRONOLOGY OF DATA RUNS

Date	Run Number	Pilot
6-6-63	None	Glenn
6-27-63	1	Sierra Sam
	2	Unmanned
6-29-63	3	
6-29-63	4	
6-29-63	5	
6-29-63	6	Sierra Sam
6-29-63	7	
6-29-63	8	
6-29-63	9	
6-29-63	10	
6-29-63	11	
6-29-63	12	
6-29-63	None	Dr. Matter
6-30-63	20	Capt. Nev. Clark
6-30-63	21	
6-30-63	22	
6-30-63	23	
6-30-63	24	
6-30-63	25	
6-30-63	26	Dr. Temple
6-30-63	27	
6-30-63	28	
6-30-63	29	
7-2-63	40	Paul LaChance
7-2-63	41	
7-2-63	42	
7-2-63	43	
7-2-63	44	Mike McMillan
7-2-63	45	
7-2-63	46	
7-2-63	47	
7-2-63	48	
7-2-63	49	Gary McNabb
7-2-63	50	
7-2-63	51	
7-2-63	52	
7-3-63	1A	Al Ciplickas
7-3-63	2A	
7-3-63	3A	
7-3-63	4A	

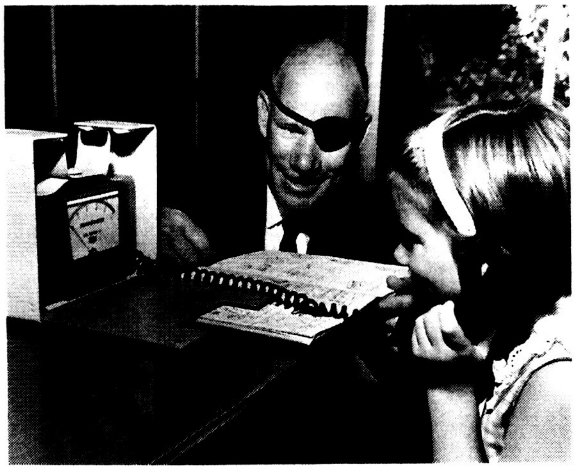
7-3-63	5B	Paul LaChance
7-3-63	6B	
7-3-63	7B	
7-3-63	8B	
7-3-63	9C	Mike McMillan
7-3-63	10C	
7-3-63	11C	
7-3-63	12C	
7-3-63	13D	Gary McNabb
7-3-63	14D	
7-3-63	16D	
7-8-63	B26	Dr. Temple
7-8-63	B27	
7-8-63	B28	
7-8-63	B30	
7-8-63	B31	Al Ciplickas
7-8-63	B32	
7-8-63	B33	
7-8-63	B34	
7-8-63	B35	
7-8-63	B36	
7-8-63	B37	Paul LaChance
7-8-63	B38	
7-8-63	B39	
7-8-63	B40	
7-8-63	B41	
7-9-63	B49	Capt. Nev. Clark
7-9-63	B50	
7-9-63	B51	
7-9-63	B52	
7-9-63	B53	
7-9-63	B54	
7-9-63	B55	Gary McNabb
7-9-63	B56	
7-9-63	B57	
7-9-63	B58	
7-9-63	B59	
7-9-63	B60	
7-9-63	B61	Mike McMillan
7-9-63	B62	
7-9-63	B63	
7-9-63	B64	
7-9-63	B65	
7-9-63	O2	
7-9-63	B73	Al Ciplickas
7-10-63	04	Major Gorman
7-10-63	05	

7-10-63	01A	Deke Slaton
7-10-63	02A	
7-10-63	03A	
7-10-63	04A	
7-10-63	05B	Major Gorman
7-10-63	06B	
7-10-63	07B	
7-10-63	08B	
7-10-63	09C	Warren North
7-10-63	010C	
7-10-63	011C	
7-10-63	012CA	
7-10-63	012B	
7-10-63	013D	John Young
7-10-63	014D	
7-10-63	015D	
7-10-63	016D	
7-10-63	B85	Dr. Temple
7-10-63	B86	
7-10-63	B87	
7-10-63	B88	
7-10-63	B89	
7-10-63	B90	
7-10-63	06	
7-11-63	019C	Jim Shows
7-11-63	020C	
7-11-63	020C2	
7-11-63	023C	
7-11-63	021B	Wally Schirra
7-11-63	022B	
7-11-63	023B	
7-11-63	024B	
7-11-63	024B2	
7-11-63	024B3	
7-11-63	024B4	
7-11-63	024B5	
7-11-63	024B6	
7-11-63	024B7	
7-11-63	025A	Gordon Cooper
7-11-63	026A	
7-11-63	027A	
7-11-63	028A	
7-11-63	028A2	
7-11-63	028A3	
7-11-63	028A4	
7-11-63	028A5	
7-11-63	028A6	
7-11-63	028A7	

ALTO SCIENTIFIC COMPANY ANNOUNCES

THE

SIGMATOMETER



A New Teaching Aid

for

The Speech and Hearing Therapists

ALTO SCIENTIFIC COMPANY, INC.
4083 TRANSPORT STREET, PALO ALTO, CALIFORNIA ▪ PHONE: 321-3434

3

A leading instrumentation advancement in the field of speech and hearing therapy has been announced by the Alto Scientific Company This new portable, all transistorized instrument is called the Sigmatometer. It is capable of distinguishing measureable differences between acceptable production of the (S) phoneme and frontal, dull or lateral lisps associated with faulty emission of this family of sounds. The instrument responds with a metered needle-and-dial measurement, providing an immediate indication of success. In field testing, teachers and clinicians have found that the Sigmatometer provides excellent motivation for improvement of speech. A child can now see how good his "S" is and compare it with that of other children. Experience has shown that the "immediate response" factor controls the attention of the student for longer periods of time than any other device now available. As a result of the attention factor, the rate of improvement in speech is accelerated – sometimes dramatically.

Suggested Applications:

- Speech defective children and adults will make significant gains in ability to discriminate fine differences in articulation of the "S" sound.
- Hard of hearing and deaf children are able to identify instantly a correct response.
- Children and adults who have cerebral palsy involvements, post operative conditions or other physical abnormalities which affect speech, may also profit from its use.

SIGMATOMETER SPECIFICATIONS

- **Physical Data:** Light-weight, rugged aluminum case (under six pounds) with baked-on scratch resistant textured finish. Physical size is 8" high, 9" wide and under 4" thick.
- **Controls:** Power Switch – On-Off; Sensitivity; and Meter Return – Fast-Slow.
- **Response Indicator:** Extra large 4" moving pointer meter display graduated by numbers and five brilliant colors.
- **Microphone:** Special high-quality microphone enclosed in a conventional telephone handset.
- **Teaching Aids:** Operating instructions and three attractive flash cards displaying selected words that have the (S) phoneme in the initial, medial, and final positions.
- **Power:** Normal household power – 115V, 60 cycles.
- **Extra Features at Additional Cost:**
  - ▲ Battery operated (Model S-1B) permits use of the Sigmatometer in locations where there are no easily available electrical outlets.
  - ▲ Detachable shoulder strap is available to free the therapists hands for carrying other equipment.

APPENDIX C

Description of Voiscon As A

Noise Tolerant Voice Switch

7-12-63	029B	Elliot See
7-12-63	030B	
7-12-63	031B	
7-12-63	032B	
7-12-63	032B2	
7-12-63	032B3	
7-12-63	032B4	
7-12-63	032B5	
7-12-63	032B6	
7-12-63	032B7	
7-12-63	033B	Tom Stafford
7-12-63	034B	
7-12-63	035B	
7-12-63	036B	
7-12-63	037B	
7-12-63	038B	
7-19-63	16A	George Cooper
7-19-63	17	
7-19-63	18	
7-19-63	19	
7-19-63	21	Bob Innis
7-19-63	22	